



数据库配置

Enterprise applications

NetApp
May 09, 2024

目录

数据库配置	1
Oracle数据库块大小	1
Oracle数据库参数: db_file_multiblock_read_count	1
Oracle数据库参数: filesystemio_options	2
Oracle Real Application Clusters (RAC)超时	2

数据库配置

Oracle数据库块大小

ONTAP在内部使用可变块大小、这意味着可以为Oracle数据库配置所需的任何块大小。但是、文件系统块大小可能会影响性能、在某些情况下、较大的重做块大小可以提高性能。

数据文件块大小

某些操作系统可选择文件系统块大小。对于支持Oracle数据文件的文件系统、使用数据压缩时、块大小应为8 KB。如果不需要数据压缩、则可以使用8 KB或4 KB的块大小。

如果将数据文件放置在具有512字节块的文件系统上、则可能会出现文件错位。LUN和文件系统可能已根据NetApp建议正确对齐、但文件I/O可能错位。这种错位会导致发生原因出现严重的性能问题。

支持重做日志的文件系统所使用的块大小必须是重做块大小的倍数。这通常要求重做日志文件和重做日志本身都使用512字节的块大小。

重做块大小

在重做率非常高的情况下、4 KB块大小的性能可能会更好、因为重做率较高、可以在更少、更高效的操作中执行I/O。如果重做速率大于50 Mbps、请考虑测试4 KB的块大小。

在块大小为4 KB的文件系统上使用块大小为512字节的重做日志和许多非常小的事务时、客户发现了一些数据库问题。对一个4 KB文件系统块应用多个512字节更改所涉及的开销导致了性能问题、这些问题通过将文件系统更改为使用512字节的块大小得以解决。



* NetApp建议*不要更改重做块大小、除非相关客户支持或专业服务组织建议您更改块大小、或者更改基于官方产品文档。

Oracle数据库参数：db_file_multiblock_read_count

。db_file_multiblock_read_count 参数用于控制Oracle在顺序I/O期间单次操作读取的Oracle数据库块的最大数量

但是、此参数不会影响Oracle在任何和所有读取操作期间读取的块数、也不会影响随机I/O只有顺序I/O的块大小会受到影响。

Oracle建议用户不要设置此参数。这样可以使数据库软件自动设置最佳值。这通常意味着、将此参数设置为一个可产生1 MB I/O大小的值。例如、读取1 MB的8 KB块需要读取128个块、因此、此参数的默认值为128。

NetApp在客户站点发现的大多数数据库性能问题都与此参数的设置不正确有关。在Oracle版本8和9中更改此值有充分的理由。因此、参数可能会在不为人所用的情况下出现在中 init.ora 文件、因为数据库已原位升级到Oracle 10及更高版本。与默认值128相比、原有设置8或16会严重损害顺序I/O性能。



* NetApp建议*设置 db_file_multiblock_read_count 参数不应出现在中 init.ora 文件NetApp从未遇到过更改此参数可提高性能的情况、但在许多情况下、它会明显损坏顺序I/O吞吐量。

Oracle数据库参数：filesystemio_options

Oracle初始化参数 filesystemio_options 控制异步和直接I/O的使用

与通常的看法相反、异步I/O和直接I/O并不互相排斥。NetApp发现、在客户环境中、此参数经常配置不当、这种配置不当直接导致许多性能问题。

异步I/O意味着可以并行处理Oracle I/O操作。在各种操作系统上提供异步I/O之前、用户配置了大量dbwriter进程并更改了服务器进程配置。使用异步I/O时、操作系统本身会代表数据库软件以高度高效的并行方式执行I/O。此过程不会使数据面临风险、Oracle重做日志记录等关键操作仍会同步执行。

直接I/O会绕过操作系统缓冲区缓存。UNIX系统上的I/O通常流经操作系统缓冲区缓存。这对于不维护内部缓存的应用程序非常有用、但Oracle在SGA中具有自己的缓冲区缓存。在几乎所有情况下、最好启用直接I/O并将服务器RAM分配给SGA、而不是依赖操作系统缓冲区缓存。Oracle SGA可以更高效地使用内存。此外、当I/O流经操作系统缓冲区时、它会受到额外处理的影响、从而增加了缓存。当低延迟是一项关键要求时、对于写入I/O负载繁重的情况、延迟增加尤为明显。

的选项 filesystemio_options 是：

- **async.** Oracle将I/O请求提交给操作系统进行处理。此过程允许Oracle执行其他工作、而不是等待I/O完成、从而提高I/O并行处理能力。
- **directio.** Oracle直接对物理文件执行I/O、而不是通过主机操作系统缓存路由I/O。
- 无。 Oracle使用同步和缓冲I/O在这种配置中、在共享和专用服务器进程之间进行选择以及dbwriter的数量更重要。
- **setall.** Oracle同时使用异步和直接I/O在几乎所有情况下、使用 setall 最佳。



。 filesystemio_options 参数在DNFS和ASM环境中无效。使用DNFS或ASM会自动导致同时使用异步和直接I/O

某些客户过去遇到过异步I/O问题、尤其是在以前的Red Hat Enterprise Linux 4 (RHEL4)版本中。互联网上的一些过时建议仍然建议避免异步IO、因为信息过时。异步I/O在所有当前操作系统上均保持稳定。如果操作系统没有已知错误、则没有理由禁用它。

如果数据库一直在使用缓冲I/O、则切换到直接I/O可能还需要更改SGA大小。禁用缓冲的I/O可消除主机操作系统缓存为数据库提供的性能优势。将RAM重新添加到SGA可修复此问题。最终结果应该是I/O性能有所提高。

尽管Oracle SGA使用RAM几乎总是优于操作系统缓冲区缓存、但可能无法确定最佳值。例如、在具有许多间歇性活动Oracle实例的数据库服务器上、最好使用SGA大小非常小的缓冲I/O。这种安排允许所有正在运行的数据库实例灵活地使用操作系统上剩余的可用RAM。这是一种非常罕见的情况、但在某些客户站点上也发现过。



* NetApp建议*设置 filesystemio_options to `setall`但请注意，在某些情况下，丢失主机缓冲区缓存可能需要增加Oracle SGA。

Oracle Real Application Clusters (RAC)超时

Oracle RAC是一款集群软件产品、它具有多种类型的内部检测信号进程、用于监控集群的运行状况。



中的信息 "MissCount" 第节介绍了有关使用网络存储的Oracle RAC环境的重要信息、在许多情况下、需要更改默认Oracle RAC设置、以确保RAC集群在网络路径更改和存储故障转移/切换操作后不会受到影响。

磁盘超时

与存储相关的主RAC参数是 `disktimeout`。此参数用于控制表决文件I/O必须完成的阈值。如果 `disktimeout` 超过此参数后、RAC节点将从集群中逐出。此参数的默认值为200。此值对于标准存储接管和回用过程应足够。

NetApp强烈建议在将RAC配置投入生产之前对其进行全面测试、因为许多因素会影响接管或恢复。除了完成存储故障转移所需的时间之外、传播链路聚合控制协议(Link Aggregate Control Protocol、LACP)更改也需要额外的时间。此外、SAN多路径软件必须检测到I/O超时、然后在备用路径上重试。如果数据库非常活跃、则在处理表决磁盘I/O之前、必须对大量I/O进行排队和重试。

如果无法执行实际的存储接管或恢复、则可以在数据库服务器上执行缆线拉拔测试来模拟这种影响。



- NetApp建议*:
- 退出 `disktimeout` 参数、默认值为200。
- 始终全面测试RAC配置。

MissCount

。 `misscount` 参数通常仅影响RAC节点之间的网络检测信号。默认值为30秒。如果网格二进制文件位于存储阵列上或操作系统启动驱动器不在本地、则此参数可能会变得很重要。这包括启动驱动器位于FC SAN上的主机、NFS启动的操作系统以及启动驱动器位于VMDK文件等虚拟化数据存储库上的主机。

如果存储接管或恢复中断了对启动驱动器的访问、则网格二进制位置或整个操作系统可能会暂时挂起。ONTAP完成存储操作以及操作系统更改路径和恢复I/O所需的时间可能会超过 `misscount` 阈值。因此、在恢复与启动LUN或网格二进制文件的连接后、节点会立即被逐出。在大多数情况下、发生逐出和后续重新启动时不会记录任何日志消息来指示重新启动的原因。并非所有配置都会受到影响、因此、请在RAC环境中测试任何SAN启动、NFS启动或基于数据存储库的主机、以便在与启动驱动器的通信中断时RAC保持稳定。

对于非本地启动驱动器或托管的非本地文件系统 `grid` 二进制文件、 `misscount` 需要更改才能匹配 `disktimeout`。如果更改了此参数、请执行进一步测试、以确定对RAC行为的任何影响、例如节点故障转移时间。



- NetApp建议*:
- 离开 `misscount` 参数、默认值为30、除非满足以下条件之一:
 - `grid` 二进制文件位于网络连接驱动器上、包括NFS、iSCSI、FC和基于数据存储库的驱动器。
 - 操作系统通过SAN启动。
- 在这种情况下、请评估影响操作系统或访问的网络中断的影响 `GRID_HOME` 文件系统。在某些情况下、此类中断发生原因会使Oracle RAC守护进程发生拖延、从而可能导致出现 `misscount`` 基于的超时和逐出。超时默认为27秒、即的值 ``misscount 减号 reboottime`。在这种情况下、增加 `misscount` 至200以匹配 `disktimeout`。

版权信息

版权所有 © 2024 NetApp, Inc.。保留所有权利。中国印刷。未经版权所有者事先书面许可，本档中受版权保护的任何部分不得以任何形式或通过任何手段（图片、电子或机械方式，包括影印、录音、录像或存储在电子检索系统中）进行复制。

从受版权保护的 NetApp 资料派生的软件受以下许可和免责声明的约束：

本软件由 NetApp 按“原样”提供，不含任何明示或暗示担保，包括但不限于适销性以及针对特定用途的适用性的隐含担保，特此声明不承担任何责任。在任何情况下，对于因使用本软件而以任何方式造成的任何直接性、间接性、偶然性、特殊性、惩罚性或后果性损失（包括但不限于购买替代商品或服务；使用、数据或利润方面的损失；或者业务中断），无论原因如何以及基于何种责任理论，无论出于合同、严格责任或侵权行为（包括疏忽或其他行为），NetApp 均不承担责任，即使已被告知存在上述损失的可能性。

NetApp 保留在不另行通知的情况下随时对本文档所述的任何产品进行更改的权利。除非 NetApp 以书面形式明确同意，否则 NetApp 不承担因使用本文档所述产品而产生的任何责任或义务。使用或购买本产品不表示获得 NetApp 的任何专利权、商标权或任何其他知识产权许可。

本手册中描述的产品可能受一项或多项美国专利、外国专利或正在申请的专利的保护。

有限权利说明：政府使用、复制或公开本文档受 DFARS 252.227-7013（2014 年 2 月）和 FAR 52.227-19（2007 年 12 月）中“技术数据权利 — 非商用”条款第 (b)(3) 条规定的限制条件的约束。

本文档中所含数据与商业产品和/或商业服务（定义见 FAR 2.101）相关，属于 NetApp, Inc. 的专有信息。根据本协议提供的所有 NetApp 技术数据和计算机软件具有商业性质，并完全由私人出资开发。美国政府对这些数据的使用权具有非排他性、全球性、受限且不可撤销的许可，该许可既不可转让，也不可再许可，但仅限在与交付数据所依据的美国政府合同有关且受合同支持的情况下使用。除本文档规定的情形外，未经 NetApp, Inc. 事先书面批准，不得使用、披露、复制、修改、操作或显示这些数据。美国政府对国防部的授权仅限于 DFARS 的第 252.227-7015(b)（2014 年 2 月）条款中明确的权利。

商标信息

NetApp、NetApp 标识和 <http://www.netapp.com/TM> 上所列的商标是 NetApp, Inc. 的商标。其他公司和产品名称可能是其各自所有者的商标。